Beyond Graphs: Capturing Groups in Networks

R. Ramanathan[*], A. Bar-Noy[†], P. Basu[*], M. Johnson[‡], W. Ren[§], A. Swami[¶] and Q. Zhao[§]

[*]Raytheon BBN Technologies
[†]City University of New York
[‡]Pennsylvania State University
[§]University of California, Davis
[¶]Army Research Laboratory

Abstract—**Currently, the de facto representational choice for networks is graphs. A graph captures *pairwise* relationships (edges) between entities (vertices) in a network. Network science, however, is replete with *group* relationships that are more than the sum of the pairwise relationships. For example, collaborative teams, wireless broadcast, insurgent cells, coalitions all contain unique group dynamics that need to be captured in their respective networks.**

We propose the use of the *(abstract) simplicial complex* to model groups in networks. We show that a number of problems within social and communications networks such as network-wide broadcast and collaborative teams can be elegantly captured using simplicial complexes in a way that is not possible with graphs. We formulate combinatorial optimization problems in these areas in a simplicial setting and illustrate the applicability of topological concepts such as "Betti numbers" in structural analysis. As an illustrative case study, we present an analysis of a real-world collaboration network, namely the ARL NS-CTA network of researchers and tasks.

I. INTRODUCTION

The analysis of communication, social, information, economic and several other types of networks is almost always based on graphs as the basic mathematical abstraction. A (directed) graph $G = (V, E)$ is essentially a set of (ordered) cardinality-2 subsets (E) on a given set (V). This restriction to *pairwise* relationships renders graphs unable to capture higher-order interactions that are distinct from the union of pairwise interactions. In particular, the notion of a *group as a fundamental manipulable entity* is missing in current network science.

At the same time, groups occur fairly commonly in many of these networks. For example, collaborative teams, wireless broadcast, insurgent cells, coalitions all contain unique group phenomena that need to be captured in their respective networks. Over the last decade, we have seen the emergence of social media such as Facebook, blogs etc., topic based grouping of information (e.g. Wikipedia) and smart phones that facilitate group interaction. Given these trends, we contend that network science will need to look beyond graphs for a suitably general representation.

In this paper, we investigate the modeling of networks using the *(abstract) simplicial complex*. A simplicial complex on a set V is a family of arbitrary-cardinality subsets of V closed under the subset operation, that is, if a set s is

in the family, all subsets of s are also in the family. An element of the family is called a *simplex* or *face*. Figure 1 illustrates the simplicial complex on three friends A, B, and C in two possible behaviors: in one, they can only talk pairwise on the phone (left), and in the other, they can both talk pairwise and as a group (right). The group interaction is shown as a *shaded* triangle representing the simplex or face. Note that the distinction between the two situations is not possible with graphs as a model, since (A,B,C) is not allowed. Moreover, with simplicial complexes, attributes or *weights* (such as frequency of interaction, time etc.) can be attached not only to vertices and edges but also to the higher dimensional faces, which, as we shall show in later sections, is useful for many network problems.
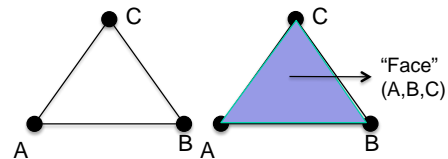

Fig. 1. The simplical complex can distinguish between three pairwise relations (left) and (additionally) a *group* relationship (A,B,C) (right).

The above example may be applied to other contexts as well. For instance, in a wireless ad hoc network, it is not possible to discern by only looking at the graph in Figure 1 (left) whether A, B, and C can communicate simultaneously over a shared broadcast channel or if they have directional antennas and so each can only talk to one other node at a time.

A natural question is: *since a group decomposable into a set of edges, why are graphs not sufficient?* It is not sufficient when there are attributes or properties of a group that are *over and above* the union of binary interactions. Note that the graph-theoretic notion of a "clique" only captures the union of pairwise relationships, and not the higher-order aggregation as the above examples illustrate. In some cases, the use of higher-order aggregations provides representation and manipulation convenience (e.g. assigning cost vector to an entire face) and in some it brings forth new structural features (e.g. "cavities" that we discuss later).

Simplicial complexes are well established in mathematics, in particular algebraic topology [1], [2], and a rich body of deep results exist on their properties. Applications to image

Report Documentation Page		*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE **APR 2011**	2. REPORT TYPE	3. DATES COVERED **00-00-2011 to 00-00-2011**
4. TITLE AND SUBTITLE **Beyond Graphs: Capturing Groups in Networks**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Army Research Laboratory,2800 Powder Mill RD,Adelphi,MD,20783-**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Presented at the Third IEEE International Workshop on Network Science For Communication Networks, April 15, 2011 , Shanghai, China

14. ABSTRACT
Currently, the de facto representational choice for networks is graphs. A graph captures pairwise relationships (edges) between entities (vertices) in a network. Network science however, is replete with group relationships that are more than the sum of the pairwise relationships. For example, collaborative teams, wireless broadcast, insurgent cells, coalitions all contain unique group dynamics that need to be captured in their respective networks. We propose the use of the (abstract) simplicial complex to model groups in networks. We show that a number of problems within social and communications networks such as networkwide broadcast and collaborative teams can be elegantly captured using simplicial complexes in a way that is not possible with graphs. We formulate combinatorial optimization problems in these areas in a simplicial setting and illustrate the applicability of topological concepts such as ?Betti numbers? in structural analysis. As an illustrative case study, we present an analysis of a real-world collaboration network, namely the ARL NS-CTA network of researchers and tasks.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**7**	

processing have been fairly well studied (see for example [3]). In communications networks, simplicial complexes have been used for studying sensor network coverage[4], [5], and the analysis of contact times in a mobile ad hoc network [6]. The problems studied in these papers are quite different from our problem, which is to capture groups. There has been some consideration of simplicial complexes as part of a mathematical framework called Q-analysis to analyze general structures [7], which some have applied to specific social network problems [8], [9]. We are not aware of work that uses simplicial complexes to model group phenomena in problems across communication, social and information networks.

Why not use "hypergraphs"[10], which also allows arbitrary-cardinality subsets? A key difference is that, unlike a simplicial complex, the subsets (hyperedges) in a hypergraph do not need to be closed under the subset operation. For example, given vertices A, B, C, the set $\{(A, B), (A, B, C)\}$ is a hypergraph, but it is not a simplicial complex because of the absence of (B, C) and (A, C). Many of the phenomena in network science that we have encountered do exhibit subset closure (e.g. friend group, broadcasting, collaboration, etc.), and therefore we believe simplicial complexes are a better fit. Further, with the closure restriction removed, some interesting and useful properties such as "cavities" are not applicable. Finally, hypergraphs do not allow succinct representation that simplicial complexes can provide by just describing the highest dimensional faces. These points notwithstanding, we believe that hypergraphs also merit investigation since results with simplicial complexes are not applicable to groups not closed under the subset operation[1]. However, for the reasons mentioned above, we have chosen to first study the applicability of simplicial complexes, leaving hypergraphs as a topic for future work.

There are a number of problems across different network types for which simplicial complexes offer an elegant abstraction. Due to space restrictions, we shall focus on two representative problems, one in communications networks and one in social networks and describe the applicability of simplicial complexes to those in detail. Specifically, in section III, we discuss *network-wide broadcast*, that is, sending a packet from a source to all nodes in a multihop wireless network. In particular, we show how weighted versions of a *neighborhood complex* provides a model that captures the group aspect of real-world broadcasting better than conventional graph-based models.

In section IV, we discuss the higher-dimensional analysis of structure and information flow in *collaboration networks*. We show how concepts such as "Betti numbers" and higher dimensional "cavities" can provide insights not possible with graphs. In section V, we briefly list a number of other problems in communications, social and information networks for which a simplicial complex will be useful. Along the way, we

formally state a number of optimization problems as possible near-term research pursuits.

Finally, as a case study, we present in section VI a simplicial model of a real-world data set. This data set is the network of collaborations within the Network Science Collaborative Technology Alliance (CTA) program of which this work is a part. Using metrics unique to a simplicial model, we analyze the various parts of the collaboration as well as the entire network.

It is not our intention to propose simplicial complexes as a generalized replacement for graphs, but simply as an additional tool to be used when higher-order group dynamics need to be captured. In the rest of this paper, we hope to convince the reader that there are several such situations, and in these situations, the use of simplicial complexes have the potential to provide new insights not possible with graphs.

II. THE SIMPLICIAL COMPLEX

An *abstract simplicial complex* (ASC) is denoted by $\Delta = (V, S)$ where V is a set of vertices, S is a non-empty set of subsets (*simplices*) of V closed under the subset operation (that is, for any $S_k \in S$, all subsets of S_k are also in S). Every abstract simplicial complex has a geometric realization as a (non-abstract) *simplicial complex* in a space of sufficient dimension. This correspondence is helpful in visualization, that is, one can think of an (abstract) simplicial complex as lines, triangles, tetrahedra and so on "glued" together in space. For the remainder of this document, we shall use "simplicial complex (SC)" synonymously with "abstract simplicial complex."

A *simplex* or a *face* of an SC $\Delta = (V, S)$ is any subset $s \in S$. The *dimension* of a simplex is one less than the number of vertices in it. The dimension of a simplicial complex is the maximum dimension of the simplices in it. A graph is a special case of a simplicial complex, i.e., an SC of dimension 1. A *facet* of a complex is a maximal face, that is, a face that is not a subset of any other face. The *i-skeleton* of a simplicial complex is the collection of all its faces of dimension $\leq i$.
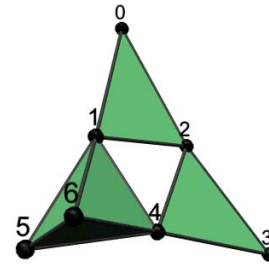

Fig. 2. An example simplicial complex

Figure 2 shows an example simplicial complex. The facets are (0,1,2), (2,3,4), and (1,4,5,6), and the faces (simplices) are all subsets of the facets, and the facets themselves. Note that (1,2,4) is not a face even though (1,2),(1,4) and (2,4) are faces. The dimension of this simplicial complex is 3. The 1-skeleton

---

[1]A rough analogy may be made to undirected vs directed graphs. While directed graphs are more general, undirected graphs are more popular as the tighter fit for most applications. On the other hand, undirected graphs cannot model all relationships.

is its underlying graph, that is, the union of all edges and vertices (no faces).

A *weighted simplicial complex* (WSC) $\Delta = (V, S, w)$ where V is a set of vertices, S is a closed set of subsets of V, and $w : S \to \Re$ is a weight function. We have found very little work on WSCs from the mathematical community, but we have found that they nicely model many optimization problems in communication and social networks.

The concept of *Betti numbers*[2] can be used to distinguish topological spaces. Intuitively, the kth Betti number B_k is the number of unconnected (via higher dimensions) k-dimensional surfaces. Specifically, B_0 is the number of connected components, B_1 is the number of 2-dimensional "holes", and B_2 is the number of 3-dimensional voids and so on. The simplicial complex in Figure 2 has $B_0 = 1$, $B_1 = 1$ (the hole (1,2,4) in the middle), and $B_2 = 0$ (no voids).

We have only given the bare minimum background required for understanding the rest of the paper. Readers interested in learning more about simplicial complexes, Betti numbers and algebraic topology in general are referred to [1], [2].

III. Broadcasting in a Multi-hop Wireless Network

The broadcast nature of the wireless channel results in a natural grouping of nodes based on the relation "the set of nodes that receive a packet via a given transmission". This is clearly closed under the subset operation (if a set of nodes receive a packet, any subset does so as well) and therefore a set of such "broadcast domains" can be aptly modeled as a simplicial complex.

In a *multihop wireless network* (MWN)[3], it is often necessary to do a *network-wide broadcast*, that is, send a packet from a given source to all nodes in the network multi-hopping through intermediate nodes. Examples include clock synchronization messages or routing control messages [11], [12]. The network-wide broadcast problem is to determine, at each hop in the sequence of broadcasts, the set of recipients that should re-transmit the packet so that the packet reaches all nodes in the most efficient manner.

Traditionally, this has been modeled using graphs as the *minimum connected dominating set*[4] problem[13], [14]. This, however, does not capture many real-world needs for several reasons. First, if the transmission needs to be reliable, that is, acknowledged, then the cost of the tree depends upon the number of *receivers* as well. Second, in rate-adaptive networks, transmissions need to use the lowest rate (highest range) that can reach the furthest receiver, and therefore each subset of receivers incurs a different cost. Third, if directional antennas are used, a dominator does not reach all its neighbors in a single transmission. Thus, what we need is a representation

---

[2]The name was coined by Henri Poincare after the Italian mathematician Enrico Betti

[3]An MWN is a peer-to-peer infrastructure-less network architecture of possibly mobile nodes which communicate over multiple hops. Examples include mobile ad hoc networks, sensor networks, mesh networks etc.

[4]A dominating set of nodes is one in which every node is either in the set or a neighbor of a node in the set.

---

in which each subset of possible receivers is a separate entity, and is associated with a possibly different cost. A simplicial complex is a natural fit for this need.

We apply the concept of a *neighborhood complex*, invented by Lovasz [15]. The *neighborhood complex* of a graph G, denoted by $\mathcal{N}(G)$ is the set of simplices such that all vertices in a given simplex share a common neighbor. In our case, we adapt the notion to mean that nodes in a simplex simultaneously receive a given transmission, that is, form a broadcast domain. Note that in some cases, such as directional transmissions, this may be different from the set of possible neighbors. Since any subset of such receivers also receive the transmission at the same time, a neighborhood complex defined thus meets the requirement of being closed under the subset operation. Figure 3 shows a graph (left) and its neighborhood complex (right), with simplices labeled by the common vertex for that simplex (ignore the rightmost figure for now). Textually, we shall represent a simplex of a neighborhood complex as $(s)[m]$ where $s \in \mathcal{N}(G)$ is a simplex and m is the list of common vertices.
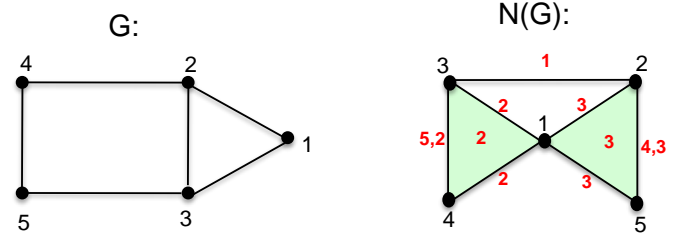

Fig. 3. A graph (G) and its labeled neighborhood complex N(G)

A broadcast transmission from a node p can be targeted to any subset s of the neighbors of p. Equivalently, in the neighborhood complex, such a transmission "activates" the simplex corresponding to s and all simplices in its closure – basically, all simplices labeled p of cardinality less than or equal to $|s|$. Thus a network-wide broadcast sequence corresponds to a sequence of simplex "activations" in the neighborhood complex. Such a sequence obviously needs to be connected. In our model, the neighborhood simplices are linked by the label, that is, the label of a simplex must be one of the vertices in the simplex of a prior simplex activation in the sequence. For example, in Figure 3, a solution to the network-wide broadcast from node 1 would be (2,3)[1], (1,3,4)[2], (2,5)[4] – notice that the label in every step i is part of a simplex in some step $j < i$ so the sequence is connected.

An alternate, and algorithmically more convenient way to do the above is to define an auxiliary graph H according to the conditions above and simply ask for a set of simplices that induce a connected subgraph in H. We state the problem below.

PROBLEM 3.1: Minimum Connected Neighborhood Cover (MCNC): Let G be the communication graph of an MWN. Let $\mathcal{N}(G)$ denote the neighborhood complex of G. Let $w : S \to \Re^+$ be a cost function on simplices $S \subseteq N(G)$. Define

an auxiliary graph H in which the vertices correspond to simplices in N(G) and there is an edge from vertex u to vertex v iff there exists a vertex $w \in S(u)$ for which $S(v)$ is a subset of the neighbors of w in G ($S(x)$ is the simplex corresponding to x). Find the set of simplices in $\mathcal{N}(G)$ of minimum total cost that induces a connected graph in H.

What is the advantage of the above approach over the relatively simple graph-theoretic formulation of minimum connected dominating set? The power of the above lies in the fact that an MCNC solution is likely to provide better broadcast distribution as it is cost-aware. The cost function can be defined based on the protocol characteristics and user preference to apply to a variety of networks. Denoting by α the cost of transmission, and by β the cost per receiver (considering, for example ACKs, but could also include receive energy), a few examples of cost functions are

- $w(S) = \alpha + \beta \cdot |S|$. This models reliable link-level multicast where one broadcast is followed by ACKs from each intended receiver.
- $w(S) = \beta \cdot |S|$ if $|S| \leq T_1$, else $k \cdot \alpha$. This models doing unicast if the number of neighbors is less than a threshold, else k link-level broadcasts. This models operations in the DARPA WNaN network [16].

We note that the number of simplices in a neighborhood complex could be exponential. However, the subset closure property allows us to maintain only facets, which is far fewer. In the weighted context, not all weight functions have polynomial-time memory, but the three discussed above do.

In sum, a generalized, cost-aware version of the network-wide broadcast problem can be captured by a simplicial complex that has, for each set of neighbor recipients, a weighted neighborhood simplex representing the cost of sending to that simplex. This is not possible in graphs which only have weightable *edges*. Thus, using simplicial complexes, a solution for problem 3.1 yields a general, cost-sensitive, realistic network-wide broadcast algorithm.

IV. COLLABORATION NETWORKS

A collaboration network is a pool of people organized into a set of teams/tasks, with each collaboration working toward a goal that helps the overall mission. Examples include collaborative research centers, task teams within a company, co-authorship for publications etc. Collaboration networks can also emerge organically, for instance the network of blogs of researchers in a specialized area.

Collaboration networks have traditionally been analyzed using graphs ([17], [18]). However, graphs cannot distinguish between different "orders" of collaboration. With graphs, three separate collaborations A-B, B-C, C-A are represented the same as a collaboration A-B-C, whereas in reality they are much different. Simplicial complexes offer a natural way to capture this distinction, and bring out interesting features.

The representation of a collaboration network as a simplicial complex is straightforward – vertices represent people, and each collaboration of k people is a simplex of dimension $k-1$. A person can be in multiple simplices. Figure 2 in section II

models a collaboration network of 7 people organized into three overlapping teams of 3, 3, and 4 members.

To illustrate the application of simplicial complexes, we consider two example problems in collaborative social networks. First, consider the question: are there potentially useful collaborations that appear to have been missed? That is, people who appear to be "near" each other in terms of interests, but are not collaborating directly. For instance, in Figure 2 in section II, the collaboration (1,2,4) is "missed". Some of these missed collaborations can be identified by "cavities" in the topological space of the simplicial complex. The existence of cavities is an application of Betti numbers as defined in section II. For example, the first Betti number identifies the number of 2-dimensional collaborations that are absent when each possible 1-dimensional sub-collaboration is present, the second Betti number identifies the number of 3-dimensional collaborations that are absent although each possible subset of 2-dimensional collaborations are present and so on.

We note that depending upon how one defines missed collaborations, there may or may not be 1-1 correspondence between cavities and missed collaborations. At the very least however, the use of Betti numbers identifies gaps in collaborations that may merit further scrutiny, especially since tools for computing Betti numbers are readily available. In section VI, we shall consider a real-life example of this.

Second, consider information flow in such a collaborative network. As information (results, event reports, opinions, rumors) flows through a social network, it is modified by the interactions of the people along the propagation path. For example, suppose a new astounding theoretical result is generated by someone in a blog collaboration network. As it disseminates through the network, it is examined and discussed by groups who scrutinize the result. Often, it is not possible for a single researcher to identify a problem but the back-and-forth discussion in a network of overlapping groups might uncover a problem or validate the result [5]. Further, it is reasonable to assume that, up to a point, the larger the group is, the more credible is the information coming out at the "other end".

This idea is captured by the concept of a *p-dimensional path*. A p-dimensional path is a connected sequence of simplices each of which has a dimension at least p. As an example, consider the network in Figure 4. The path dimension from 0 to 5 is 1 and from 0 to 8 is 2 – we argue that the information flow from vertex 0 to 8 is in some sense more "robust" than from 0 to 5 since it flows through larger groups.

It might be desirable to augment the collaboration network to achieve a certain minimum dimensionality of all paths. This leads to the following optimization problem.

PROBLEM 4.1: Dimension Augmentation: Given a simplicial complex $\Delta = (V, S)$, and a dimension requirement P, find the minimum number of faces to add so that between every pair of vertices there is at least one P-dimensional path.

---

[5] A case in point is the recent $P \neq NP$ proof attempt from Deolalikar which was discussed in the blog network, and problems identified in few days – if there was only email, it would have likely taken much longer.
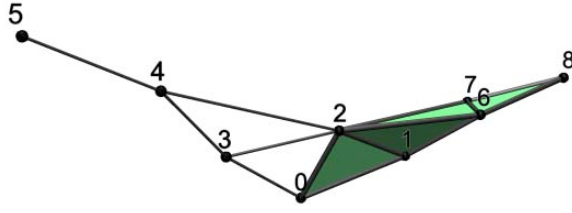
Fig. 4.   An example weighted simplicial complex to illustrate path dimension.

V.   OTHER PROBLEMS/NETWORKS

There are a number of other problem domains in which groups arise naturally and benefit from a simplicial model. Consider the problem of *team selection* from among a pool of people. Each team can be represented as a weighted simplex with a benefit function representing how well the individuals within the team work with each other. The best set of teams is then the maximum-aggregate-benefit simplicial cover. Such a problem of team selection occurs also within communications networks– *cooperative sensing* requires a team of nodes to jointly sense portions of the spectrum so as to aid dynamic spectrum access [19]. In this case the benefit may be a function of mutual distances between nodes. Another domain is cascaded *cooperative diversity* [20] in which the set of nodes that coooperatively transmit a packet needs to be selected.

The neighborhood complex introduced in section III is suitable for capturing groups that may not mutually interact but interact through a "hub". Networks based on *social media* such as Facebook and Twitter offer great examples – for instance, the set of all individuals subscribed to a Twitter feed is a neighborhood simplex. The question of how long it takes for a piece of information (or rumor) to propagate through a network is roughly similar to the network-wide broadcast problem discussed in section III. Analogs of the collaboration problem discussed in section IV occur in other domains as well – for example, the problem of *target tracking* using collaborating sensors [21], with collaborations forming faces.

Networks other than social and communication present interesting group aspects as well. In *information networks*, tightly inter-related documents or topics form groups. In [22], document clustering has been modeled in a specific way as a simplicial complex, but several variants are possible and need to be explored. The set of citations in a paper form a neighborhood complex that can be navigated much like the broadcast problem. Economic networks, political alliances, financial networks are other areas with possible group aspects.

VI.   CASE STUDY: THE NS-CTA NETWORK

In section IV, we discussed the structure of collaboration networks. In this section, we briefly study a real collaboration network, namely the Network Science Collaborative Technology Alliance (NS-CTA) network. Coordinated by the Army Research Laboratory (ARL), the NS-CTA is a collaborative network of about 80 researchers from the fields

of communications, information and social networks, and organized into six "centers", each of which has a research agenda within the broad goal of advancing network science. Each center has a number of *tasks*, each task comprising 3-7 researchers targeting a specific topic of research within the center's agenda. A researcher can be (and typically is) in more than one task[6].

Figure 5 shows the simplicial complex representation of one of the centers. The vertices are researchers and each task is a simplex (face). We used a tool called Polymake [23] for visualizing and analyzing this network. Polymake depicts a face of dimension k, representing a task with $k+1$ members, as a polyhedron of dimension k projected on to a plane.
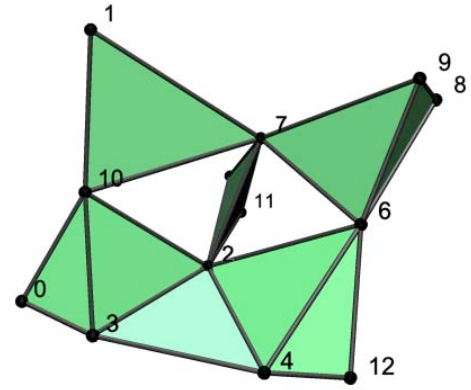

Fig. 5.   Collaboration simplicial complex of a center in the NS-CTA.

We consider the problem of "missed collaborations" discussed in section IV. As discussed there, Betti numbers can be used to identify some missed collaborations and can be readily obtained using Polymake. We shall only mention the first three Betti numbers – the others are trivial (0). The center shown in figure 5 has the Betti number sequence (1,2,0). That is, it is connected and has two 2-dimensional holes. The identification of these holes can be made visually in case of such small networks. For the center in Figure 5, the missed collaborations are 2,7,6 and 2,7,10. That is, researchers 2,7, and 6 appear close in their interests and are in pairwise collaborations, but are missing out on the fruits of 3-way group interaction. For larger networks, and higher-order cavity identification, we will need computational homology techniques [24] and tools which we are currently investigating.

We have analyzed the other five centers as well. Of the six centers, three have Betti numbers (1,2,0), that is, each of them is connected and has 2 holes. The other three have Betti numbers (2,0,0), that is, each of them is disconnected into two components with no holes. Moreover, we found that the smaller of the two components is of cardinality 3 in all three cases. As mentioned earlier, one of the uses of Betti numbers is to distinguish between topological spaces. Applying this to the data above, and thinking of the network as a topological space,

---

[6]Indeed, in a self-referential way, this paper itself stems from one such task in the NS-CTA!

it is interesting that it can be partitioned into two "classes" with great topological similarity within each class. This is all the more remarkable because the NS-CTA collaborations were formed "organically" without any central authority, and there was no particular difference in the rules concerning team formation across centers. Closer analysis is needed to determine if the topologial similarity of the three centers is the result of some underlying phenomena or mere coincidence.

We have also analyzed the entire NS-CTA network consisting of about 80 researchers, ignoring center boundaries. The NS-CTA simplicial complex has Betti numbers of (1,18,0), that is, it is connected and has 18 holes. Clearly, in each of the disconnected triples in the three centers mentioned above, there is at least one member who is also in another center, resulting in elimination of the partitions. It also appears that these centers "touch" each other at several places forming the numerous inter-center holes. A "hole" is not a "deficiency" – indeed, to make a hole the 1-skeleton (underlying graph) needs to be sufficiently well connected. Lack of holes may indicate a poorly connected network (especially if the first Betti number exceeds 1), or sufficient 2-D simplices to fill the holes.

Our investigation has shown that a simplicial model allows new and unique structural properties such as cavities and topological similarity. Since graphs do not aggregate at higher dimensions, these properties are beyond the scope of graphs. Further work is needed, however, to fully understand the applicability and adapt these tools to real-world questions.

VII. CONCLUDING REMARKS

The National Research Council (NRC) defines Network Science as "the study of network representations of physical, biological, and social phenomena leading to predictive models of these phenomena". Given the centrality of "representations" in this definition and hence the overall endeavour, it is important that we pick the right representation early on. We have argued that we need to look beyond graphs for the right representation. In particular, we have proposed the (abstract) simplicial complex as an appropriate generalization to capture group phenomena. We have illustrated several domains in communications and social networks in which a simplicial complex can provide analytical insights not easily possible with graphs.

Two broad research directions are possible in applying simplicial complexes to network science. First, combinatorial optimization problems that were based on graphs can now be re-framed in the simplicial context, along the lines of problem statements given in sections III and IV. Second, we can bring to bear results and techniques in the field of simplicial complexes, computational homology and in general algebraic topology to analysis of group phenomena in networks. We have already seen the insights from Betti numbers – but there likely are many other concepts that are applicable.

As part of the NS-CTA program, we have just begun investigating these topics. Apart from the work discussed here, we have NP-hardness results and initial approximation algorithms for some of the problems. However, these are but a small fraction of the open research problems and promising directions in this area that we believe that the network science community should investigate.

REFERENCES

[1] J. R. Munkres, *Elements of Algebraic Topology*. Addison-Wesley, 1984.
[2] E. H. Spanier, *Algebraic Topology*. Springer, 1994.
[3] W. Matusik, M. Zwicker, and F. Durand, "Texture design using a simplicial complex of morphable textures," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 787–794, 2005.
[4] A. Tahbaz-Salehi and A. Jadbabaie, "Distributed coverage verification in sensor networks without location information," *IEEE Transactions on Automatic Control*, vol. 55, pp. 1837 –1849, Aug 2010.
[5] R. Ghrist and A. Muhammad, "Coverage and hole-detection in sensor networks via homology," in *IPSN*, pp. 254–260, 2005.
[6] B. D. Walker, "Using persistent homology to recover spatial information from encounter traces," in *MobiHoc*, pp. 371–380, 2008.
[7] R. H. Atkin, "From cohomology in physics to q-connectivity in social science," *Int. J. Man-Machine studies*, vol. 4, pp. 139–167, 1972.
[8] P. Gould and A. Gatrell, "A structural analysis of a game: The liverpool v machester united cup final of 1977," *Social Networks*, vol. 2, pp. 255–273, 1979-1980.
[9] L. C. Freeman, "Q-analysis and the structure of friendship networks," *Int. J. Man Machine Studies*, vol. 12, pp. 367–378.
[10] C. Berge, *Hypergraphs*. North-Holland, 1989.
[11] T. Clausen and P. Jacquet, "Optimized link state routing," in *IETF RFC 3626*, 2003.
[12] C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing," in *Proc. 2nd IEEE Workshop on Mobile Computing Systems and Applications*, pp. 90–100, 1999.
[13] R. Gandhi, A. Mishra, and S. Parthasarathy, "Minimizing broadcast latency and redundancy in ad hoc networks," *IEEE/ACM Trans. Netw.*, vol. 16, no. 4, pp. 840–851, 2008.
[14] S. Guha and S. Khuller, "Approximation algorithms for connected dominating sets," *Algorithmica*, vol. 20, pp. 374–387, 1998.
[15] L. Lovász, "Kneser's conjecture, chromatic number, and homotopy," *J. Comb. Theory, Ser. A*, vol. 25, no. 3, pp. 319–324, 1978.
[16] M. Lawlor, "Wireless to the nth degree," *AFCEA SIGNAL online*, July 2006.
[17] M. E. J. Newman, "The structure of scientific collaboration networks," *Proc. National Academy of Sciences*, vol. 98, no. 2, pp. 404–409, 2001.
[18] J. J. Ramasco, S. N. Dorogovtsev, and R. Pastor-Satorras, "Self-organization of collaboration networks," *Phys. Rev. E*, vol. 70, no. 3.
[19] Q. Zhao and A. Swami, "Spectrum sensing and identification," in *Cognitive Radio Communications and Networks: Principles and Practice, A. Wyglinski, M. Nekovee and T. Hou, Eds*, Elsevier, Dec 2009.
[20] R. Ramanathan, "Challenges: A radically new architecture for next generation mobile ad hoc networks," in *Proc. of ACM MobiCom*, (Cologne, Germany), Aug. 2005.
[21] J. Liu, J. Reich, and F. Zhao, "Collaborative in-network processing for target tracking," *EURASIP J. Appl. Signal Process.*, vol. 2003, pp. 378–391, January 2003.
[22] T. Y. Lin and I.-J. Chiang, "A simplicial complex, a hypergraph, structure in the latent semantic space of document clustering," *Int. J. Approx. Reasoning*, vol. 40, pp. 55–80, July 2005.
[23] E. Gawrilow and M. Joswig, "polymake: a framework for analyzing convex polytopes," in *Polytopes — Combinatorics and Computation* (G. Kalai and G. M. Ziegler, eds.), pp. 43–74, Birkhäuser, 2000.
[24] T. Kaczynski, K. Mischaikow, and M. Mrozek, *Computational Homology*. Springer, 2010.